The experimental error in field trials and the effect on this error of various methods of sampling.

By

Chiaki Miyake. (*Nōgaku-Tokugioshi.*)

I. Introductory remarks.

The homogenous growth of crops on a field is impossible, since numerous factors can never act in exactly the same manner upon all individuals and, they are very complex, some of them being independent others reacting upon one another. For the same reason, the mean results of a series of experiments does not always represent the truth. The magnitude of the experimental error in field trials is variable owing to nature of crops, systems of their cultivation, irregularities of the soil and weather of the season, etc.; thereby its magnitude must be culculated by ordinary methods based upon the theory of least square, from results of series of these field trials themselves.

Thus the magnitude is variable but, I believe, it is not difficult to determine the general degree of irregularity from even a few results of series of field trials.

In general, the value of an experiment depends upon the degree of confidence which can be attached for its result. So even the measure of this general degree, seems for us to be indespensable for agricultural experimenters, as necessary to the interpretation of results of field trials, especially to the knowledge of the effect on this error of various methods of sampling. For a few examples of these questions in Japan I have worked the following measurements.

II. The experimental error of field trials on barley.

In order to illustrate the magnitude of the experimental error in field trials, I have measured the irregularity of yield on the barley which was usually cultivated in 1912 on the Experimental Farm of Imperial College of Agriculture and Forestry at Morioka. Of course, the crop has been subjected to no irregular treatment throughout the field. At the time of ripening in June, that field was divided into 100 small unit plots of each are a Tsubo[1] along the furrow. The barley which was grown on these small unit plots of a hundred numbers was harvested, air dried, threshed and then weighted up separately and carefully in same manner.

From these results,[2] I have obtaind the following numerals by ordinary method.[3]

(A) Average yield of single unit plot. .... 667 Momme[4]
(B) Average deviation on the yield of unit plots. 15.3%
(C) Standard deviation on the yield of single unit plot. ... 18.0%
(D) Standard deviation on the average yield. 1.8%

Next, I have culculated the standard deviation of successive combination plots, made up by adding the neighbouring plots together. The last step was to test the standard deviation of other combination plots which was made up by adding equi-distant scattered unit plots together.

These culculating figures were represented in the Table I.

Table I.

Size of plot. (in Tsubo)	St. dev. on new plot. (in percentage)	Estimated st. dev. (in percentage)	Combination plot that
1	18.0	18.0	(unit plot itself.)
3	11.0	10.5	made up of successive 3 unit plots.
3	10.8	10.5	" " " scattering 3 " ".
5	9.6	8.1	" " " successive 5 " ".
5	8.0	8.1	" " " scattering 5 " ".
10	7.4	5.7	" " " successive 10 " ".
10	5.9	5.7	" " " scattering 10 " ".

1.) Tsubo = 35.58 Sq. ft.

2.) On the details see "Jour. of the sci. Agric. Society, No. 127, pp. 20—29, 1913; Tokio Japan."

3.) All calculation on this are referd to "An introduction to the theory of statistics by G. U. Yule."

4.) Momme = 0.1323 Ounce.

Above secured results enable us to settle following three important points.

(a) The standard deviation decreases with the size of the plot, but the velocity seems to be slackened gradually.

(b) The standard deviation of "the scatteringly adding plot" is always smaller than that of "the successively adding plot", under the equal area.

(c) The theory of least square teaches us that the magnitude of the standard deviation varies with inverse square root of the measureing "weight" which may be represented by the area in this case. If the standard deviation of the unit plot (namely 18%) be divided by the square root of the area of combination plots, the quotient shall be the estimated standard deviation which is expressed in the third column of the Table I, and which may be right theoretically, under the presumption as it is due simply to the error in random sampling.

In comparing this estimated standard deviation with the real one in the second column of the Table I, I find that the scattering addition method is more favourable than the successive addition method in point of sampling.

III. Experiment on the rice crop.

In autum 1914, I carried out a measurement which on the same principle as that in the last article, on the rice crop at the Ōhara Institute of Agriculture at Kurashiki near Okayama.

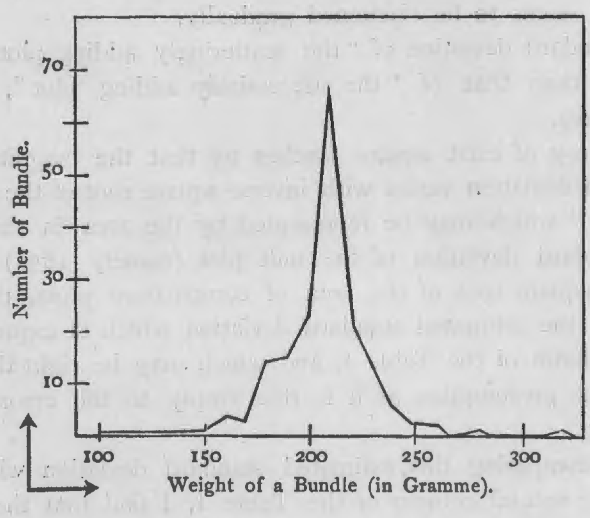
The crop and field were ordinarily and uniformly treated so that the growth of the crop did not seem apparently very irregular, except a few square feet at two corners of the field. The rice crop was planted in regular shape at the rate of 42 stubbles per Tsubo. When the crop was ready for harvesting, I reaped along the row of it rectilineally; and whenever the reaped stubbles amounted 50 in number they were tied in a bundle. Thus, I got 176 bundles by the time all crops in the field were harvested up.

These bundles were air dried for a few days, threshed, cleaned and treated to that the unhulled grain was weighted up from the first bundle to the next and next separately.

Figure I gives the frequency curve for the above secured results.[1)]

1.) On the details see "Jour. of the sci. Agric. society, No. 155, pp. 473—489, 1915; Tokio Japan".

Fig. I.



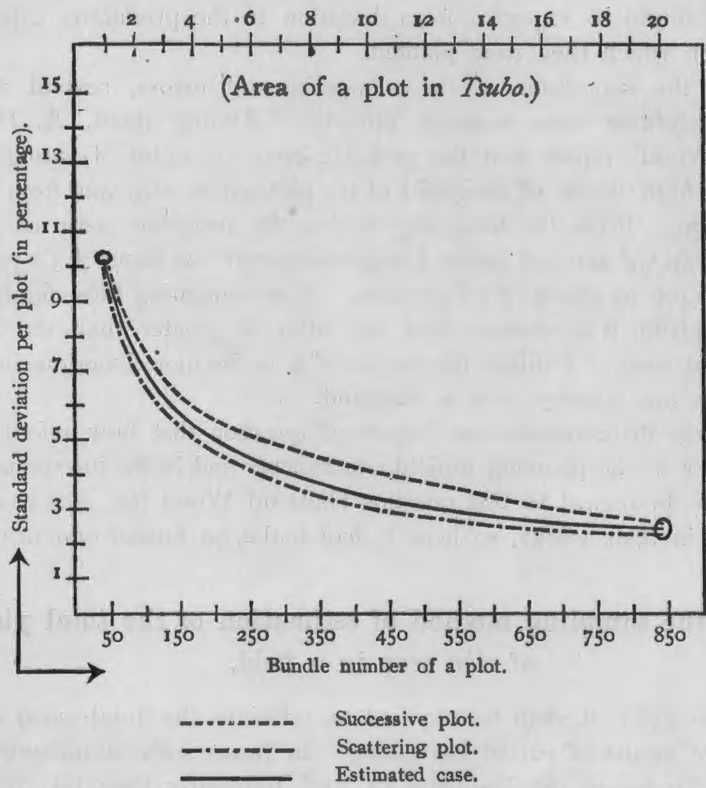
We may assume that no abnormal causes of variation are at work, and we may average our results with confidence. Furthermore, as the case of the barley, I calculated the average yields and those standard deviations on the unit plot and successive and scattering combination plots.

Table II shows these results, and Fig. II their plotted curves.

Table II.

Size of plot. (in Tsubo)	Average yield. (in Gramme)	Standard deviation. (in Percentage)	Estimated st. dev. (in Percentage)	Combination plot that
1	2090	10.3	10.3	(unit plot consisted from 50 stubbles.)
3	6257	6.4	6.0	made up of successive 3 unit plots.
3	6257	5.4	6.0	„ „ „ scattering 3 „ „.
5	10436	5.4	4.6	„ „ „ successive 5 „ „.
5	10436	4.7	4.6	„ „ „ scattering 5 „ „.
10	20879	4.0	3.3	„ „ „ successive 10 „ „.
10	20879	2.9	3.3	„ „ „ scattering 10 „ „.
17	35495	2.9	2.5	„ „ „ successive 17 „ „.
17	35495	2.4	2.5	„ „ „ scattering 17 „ „.

Fig. II.



This table furnishes the following conclusions, namely;

a) the standard deviation decreases with the size of the plot;

b) the standard deviation of "the scattering addition plot" is always smaller than that of "the successive addition plot", on an equal area;

c) the scattering addition method is always more reasonable and favourable than the successive addition method in point of sampling.

These points are the same as those of barley, except only that the standard deviation of the former was smaller than that of the latter in numerical magnitude.

IV. Interpretation of the measurements.

It is to be remarked that, such high heterogenity, as that above mentioned, exists upon the yield of crops where the whole field has been uniformly treated. Since such heterogenous yield of crops may be expected, we have some considerations upon field experimental works and upon the interpretation of their results.

Indeed, we can not conclude from a given experiment that factor A is more effective than factor B, unless the differences found are greater than those which might be expected from deviation in the productive capacity of the plot upon which they were planted.

As for the magnitude of these experimental errors, several authors have tried to obtain some measure, already. Among them, A. D. Hall and T. B. Wood[1)] report that the probable error on a plot of various crops was not far from $\pm 5\%$ of the yield of the plot whose area was from 1/2 to 1/80th. of acre. From the foregoing results, the probable error on a plot (area of 1/50th. of acre) of barley I may culculated[2)] as about ± 3.3 per cent, and of rice crop as about ± 1.8 per cent. Now comparing these figures with those in England, it is obvious that the latter is greater than the former, namely about twice. I think, the origin of it is the more laborious intensive cultivation in our country than in England.

Ultimately there remains an important question that how referd to the probable error in the planning of field experiments and in the interpretation of their results. In regard to this question Halla nd Wood (*lqc. cit.*) have fully explained it in their works, so here I shall make no further statement.

V. On the sampling method of estimation of the total yield of rice crop in a field.

In some cases, it shall be required to estimate the total yield of crop in a field by means of partial harvesting. In Japan such sampling methods are called *Tsubogari* or *Tsubogari-hō*, and frequently they take place in rice fields. However the *Tsubogari* with such purpose are objected to by many agriculturists, since it brings a great error in its results, even although it be carried out skillfully and carefully. But in fact, it leaves unknown what magnitude of error attaches to the usual method, and what system is preferable in these various methods. To solve these question, I carried out the following experiments.

Sample fields.

In the Experimental Farm of the Ōhara Institute of Agriculture, for sample fields, I chose six rice fields, the area of each of them being within one Tan[3)] or so. Each sample field was treated in the most usual way, and

1.) See "Field trials and their interpretation," by A. D. Hall, and E. J. Russel, and "The interpretation of experimental results," by T. B. Wood. On the supplement (No. 7) to the Journ. Board of Agr. Great Britain. Nov. 1911.

2.) 1/50 acre = 24 Tsubo.

$$\text{Probable error on the barley} = \pm 0.6745 \times \frac{7.4\sqrt{10}}{\sqrt{24}} = \pm 3.3$$

$$\text{Probable error on the rice crop} = \pm 0.6745 \times \frac{4.0\sqrt{10}}{\sqrt{24}} = \pm 1.8$$

3.) Tan is about $\frac{1}{4}$th. of acre.

therin 42 stubbles of the plant in each Tsubo were planted regularly in square shapes. In each of the six field, I tried partial harvesting by following five methods in same ways.

Sampling methods.

1st method. Usual method in Japan.

Let us select the most moderately grown part, at sight, in the field; and from this part take up 50 stubbles in a square shape. Then we may consider the yield of these 50 stubbles of rice plant to be representative of the whole area. Since the moderately grown part is not limited to one place only, so from another two similarly grown parts I reaped up 50 stubbles each. The first I will call 1A plot, the second 1B and the third 1C.

Furthermore the plot 1D was made up by getting together the results of plots 1A, 1B and 1C.

2nd method.

The most representatively grown part in the field was selected, and it was limited to 10 Tsubo marked with two ropes, each 12 and 30 feet long, to make a rectangle. And then, from those along the two oposite shorter sides of that rectangle and its middle line, 30 ordinary stubbles of rice plants were harvested together.

3rd method. Scattering method.

By this method, from every ten scattered places that appear moderately grown ten ordinary stubbles are cut together, and then the total numbers of sample stubbles amount to one hundred.

4th method. Diagonal method.

In the first place, let us find out the principal four corners in the field, and stretch to ropes between each two opposite corners diagonally. Next I reaped the nearest stubble to the diagonal rope, in each row of plants, and I gathered them in two diagonal groups. *Any* one of these groups I called the 4A, and the *other* the 4B and the sum of these two the 4C.

5th method. Dr. Inagaki's method.

By this method, we first estimate roughly the total number of stubbles in the field and assume the estimated numbers divided by a hundred, then we may know the stubble number N in each of hundred divisions.

In the next place, counting the total number of plants rectlinearly from the first row to the second and so on, we come at number N; then we take *any* one stubble from the first division, as the first reaped stubble. The first reaped stubble being so determined the stubble which falls on the Nth from the first shall be cut off as the second.

In the same way, the third, the fouth......and ultimately to about the hundredth stubble should be harvested scattered at equal distance.

Result.

In these ways, I had sixty plots as a whole and these results may be mention as follows:

Table III.

Sample field.	Sampling method.	Number of the cutting Stubbles.	Yield of grain. (in Momme)	The yield. (in Percentage)
E4	1A.	50	643	106.6
E4	1B.	50	685	113.6
E4	1C.	50	666	110.5
E4	1D.	150	1994	110.3
E4	2.	30	462	127.7
E4	3.	100	1182	98.0
E4	4A.	117	1328	94.0
E4	4B.	110	1194	90.0
E4	4C.	227	2522	92.1
E4	5.	92	1063	95.8
E5	1A.	50	513	94.3
E5	1B.	50	543	99.8
E5	1C.	50	556	102.2
E5	1D.	150	1612	98.7
E5	2.	30	323	98.9
E5	3.	100	1113	102.3
E5	4A.	57	598	96.4
E5	4B.	45	510	104.1
E5	4C.	102	1108	99.8
E5	5.	96	1046	100.1
E6	1A.	50	541	102.4
E6	1B.	50	547	103.6
E6	1C.	50	461	87.2
E6	1D.	150	1549	97.8
E6	2.	30	314	99.1
E6	3.	100	1091	103.3
E6	4A.	44	454	97.7
E6	4B.	40	426	100.8
E6	4C.	84	880	99.2
E6	5.	87	930	101.2

Sample field.	Sampling method.	Number of the cutting Stubbles.	Yield of grain. (in Momme)	The yield. (in Percentage)
W2	1A.	50	633	99.8
W2	1B.	50	662	104.4
W2	1C.	50	692	109.1
W2	1D.	150	1987	104.4
W2	2.	30	462	121.4
W2	3.	100	1191	93.9
W2	4A.	61	768	99.3
W2	4B.	60	781	102.6
W2	4C.	121	1549	100.9
W2	5.	105	1232	92.5
W3	1A.	50	447	97.3
W3	1B.	50	430	93.6
W3	1C.	50	426	92.7
W3	1D.	150	1303	94.5
W3	2.	30	222	80.5
W3	3.	100	936	101.8
W3	4A.	59	583	107.5
W3	4B.	58	549	103.0
W3	4C.	117	1132	105.3
W3	5.	102	993	105.9
W5	1A.	50	451	102.03
W5	1B.	50	438	99.1
W5	1C.	50	438	99.1
W5	1D.	150	1327	100.1
W5	2.	30	252	95.0
W5	3.	100	877	99.2
W5	4A.	75	660	99.5
W5	4B.	26	222	96.6
W5	4C.	101	882	98.8
W5	5.	94	862	103.7

From this table I grouped separately the similar plots of sampling, and calculate their average yields and standard errors of each method.

Table IV shows these results.

Table IV.

Sampling method.	Average yield in percentage.	Standard error on one plot.	Standard error on the average.
1A	100.4	± 3.95	±1.61
1B	102.3	± 6.14	±2.51
1C	100.0	± 8.30	±3.39
1D	101.0	± 5.10	±2.08
2	103.8	±16.05	±6.55
3	99.7	± 3.19	±1.30
4A	99.1	± 4.20	±1.72
4B	99.5	± 4.89	±2.00
4C	99.3	± 3.87	±1.58
5	99.9	± 4.55	±1.86

From this table we can summarise the following facts.

(1) The yields of those methods in which moderately grown parts were taken, as the first and second method, always seem to us to show too much yield. This may be explained from the fact that what we recognize, at a glance, to be moderate or representative in a group, is frequently better or larger than the true mean.

We experimenters must be more careful in this point!

(2) The nearest resulting plot to the average yield (namely 100%) is that of the first method, but the farthest is that of the second. However the value of the method should not be determined by results of the yield only, rather it is due to the error attached to it. From this point of view, it will be better to prefer the third or fourth, and next the fifth method. The second method seems to be a bad one.

(3) The absolute magnitude of errors attached to the various methods are as foregoing, but for the degree of accuracy there must be taken into account the number of plants, taken as a sample.

The number of cut stubbles, as in table III shown, were different according to those methods and errors of methods may vary with inverse square root of the number of those sampled plants.

Accordingly, next, I will test how these standard errors are affected by the number of plants sampled, that is to say, which method was more precisely related to the number of plants sampled. Now let n be the total sampled number of any method and σ be its real standard error. And suppose 300 stubbles were taken equally in each method, so the estimated standard error ω shall be culculated through the equation, namely

$$\omega = \sigma \times \sqrt{\frac{n}{300}}.$$

Table V shows these culculated results.

Table V.

Sampling method.	n	σ	ω
	Total sampled number.	Real standerd error.	Estimated standerderror.
1A	300	±1.61	±1.61
1B	300	±2.51	±2.51
1C	300	±3.39	±3.39
1D	900	±2.08	±3.60
2	180	±6.55	±5.07
3	600	±1.30	±1.84
4A	413	±1.72	±2.01
4B	339	±2.00	±2.13
4C	752	±1.58	±2.50
5	576	±1.86	±2.58

From this table we may conclude that, the most preferable method is the 3rd or the 4th. The first method is worse than the 3rd or the 4th, since magnitudes of the standard error varied with the plot as 1A, 1B and 1C; and the fifth is too troublesome for practical purposes.

VI. Sammary.

(1) To test the yield heterogenity of crop in a field I tried dividing the field into small plots and secured the deviation of yields on these small plots. According to these results, the probable deviation of the single plot whose area was one-fiftieth of an acre, was about ±2 per cent in terms of the yield.

(2) It is to be remarked that the result of a field experiment, which takes place in such area per plot, has about ±2 per cent as the *experimental error*.

(3) In order to estimate the total yield of a field crop by means of partial harvesting, there should be held various methods of sampling.

In those, however, we may take the diagonal or the scattering method, favourably.

(4) The single plot of any field experiment, in order that the experimental error does not increase should possess a tolerably large area.

And further, every experimental plot has not single, simply; but it is divided into several small plots and these small plots should be distributed over the field scatteringly.